A Combination of Representation Styles for the Acquirement of Speech Abilities

Virginie Govaere

A combination of representation styles for the acquirement of speech abilities

Virginie Govaere
LORIA & Universite Henry Poincare
BP 239, F- 54506 Vandoeuvre-les-nancy, France
Virginie.Govaere@loria.fr
Tel. [33](0)383 59 20 74

Abstract:
In this article, we present GEREV, a guidance and assessment system for voice rehabilitation, and in particular, a system for representation and structuring of manipulated knowledge. The first point that we develop concerns details of the knowledge domain. Indeed, contrary to many Intelligent Tutoring Systems (ITS) systems which allow acquisition of reasoning (for example in geometry or in declarative knowledge), we work on the acquisition of the norm of acoustic parameters; these can be defined as an ensemble of physical sizes, some of which constitute the standard intervals of production for categories such as male or female. In rehabilitation, the norms must needs be adapted to students' possibilities. Therefore, we have centered a part of our work on the style and structuring of knowledge, which will permit us to adapt the norm.
Next, we present our choice of representations and their implications. We qualify the quantitative values with the help of fuzzy sets, which allow us on the one hand to pass from the quantitative to the qualitative, and on the other hand to obtain as exhaustive and precise a representation as possible.
At the level of the knowledge unit itself - that is the ability, defined as the entire body of information that allows to define an acoustic parameter representation - we use frames, which precisely describe notions of knowledge prototypes in the domain. As for the global organization of expertise, we chose to describe it with the help of a network of frames which express relations between abilities. These combinations of representation styles are necessary in order to express the totality of knowledge details. However, they entail modifications within modules of the expert and the student which we will describe, before presenting in conclusion, an outline of thoughts on the portability of our work on other systems; and finally, perspectives which are evoked by problems concerning implantation of such a system, student assessment, and correction and guidance in GEREV.

Introduction

The theoretical objective of this work is to model a student, with the aim of better adapting the interactive training environment placed at learner disposal. In order to individualize a system, it must arrange, notably, information on the user. With this view, we turned to the different types of knowledge used by the training system; their form and their organization as well as their utilization in guidance. Here, the originality of our contribution is at the level of the type of knowledge manipulated on the one hand, and on the other hand, at the level of its organization. In this article, we first present GEREV, a system of guidance and assessment of learning in the framework of the speech rehabilitation system SIRENE; we then discuss features of representations, which will lead us to an exposition of our representational choices. Then, we will consider propositions for the GEREV's architecture, and we will conclude with reflections on the portability of the theoretical and practical contributions of this work. Finally, in the last section, we will evoke perspectives of GEREV which bear on the problems of establishing such a system, and of assessing, correcting, and guiding students in GEREV.

Objectives and presentation of the GEREV system

GEREV is an in-progress software developed in an existing environment, SIRENE [Haton98]. Its function is to suggest an assessment and guidance adapted to the student. Distinctive features of GEREV include, on the one hand, the aim of training of a norm and not of a knowledge set, and on the other hand, the possibility to modify this norm according to the evolving ability of the student. Our objective is for the learner to acquire intelligible speech, which can be defined with the help of a particular set of abilities. What we refer to as ability throughout this article, concerns the entire data on an acoustic parameter as intensity, pitch. We do not ask the student to produce a value of an acoustic parameter within a given range, but rather to tend towards a norm while accounting for his or her deficits. By this last point, our approach distinguishes itself amongst the many ITS which apply to knowledge training, such as the training of languages, a geometry, or a research on disabilities.

The context

SIRENE is a software intended for deaf adults. It presents a graphic visualization of several parameters of the speech, such as the resonant intensity, the fundamental frequency, the rhythm of statements, the components of articulation. Normally, the speech is controlled via auditory feedback. In the case of the deaf people, this feedback can not occur. The principle of our software is relatively simple: it affords a visualisation of the acoustic parameters necessary in establishing and mastering intelligibility of speech. These parameters are regrouped in three large categories: voice, articulation, and prosody. In the present state of the system, SIRENE relies on the intervention of the speech therapist/rehabilitator, who will calibrate the exercises according to his or her level of requirement and the abilities of the subject. Furthermore, the assessment and advice are general and, in fact, must to be adapted to the student.

Therefore, if one wants to permit regular training by an autonomous system, it is necessary to attack the problem (currently not treated in the software) of generating advice and guidance adapted to the student.

From this perspective, we propose an autonomous system. It will contain, notably, knowledge on the student and on the reference's domain. The speech therapist will parameterise the system while accounting for data on the student, such as the sex, the degree of deficiency in the different groups of parameters, etc. According to these data and the current observations, the software will propose certain exercises, advices, and will be able to propose a different level of requirement according to abilities of the subject. This will bring to the level of domain knowledge, by an adaptation of the norm.

Features of manipulated knowledge

As we have previously seen, the manipulated knowledge concerns speech. The specialist expresses domain's knowledges with the help of qualitative statements of the form "the intensity is weak". The student's speech is controlled and measured rationally, with a certain number of acoustic parameters (pitch, voicing, intensity...). The quantitative measures constitute a first stage of treatment, since these raw data are not directly usable - each numeric value of a parameter does not itself suggest a qualitative assessment. It is therefore necessary to transform these quantitative data to qualitative data. For that, it is necessary to form the numeric value groups which correspond to the qualitative data. Nevertheless, it is suitable to utilise strict ranges of parameter values. Indeed, we must be able to keep account of the proximity of those values; it is quite inconceivable that two near values, such as 40 dB and 41 dB, should be categorized as belonging to two qualitatively different groups of data. It is thus necessary to have recourse to a style of knowledge representation which permits to transform from numeric to qualitative while accounting for the proximity of the numeric values; and all that, in order to be able to use a faithful picture of the subject's production and to acquire sufficient informations so as to account for the evolution of their abilities. However, our objective is not to collect the maximum amount of data on the student, but rather to build up sufficient informations to which the system will then refer for individualizing the guidance of the learning, and the updating of the norm initially aimed at. Effectively, those data must be reliable and pertinent for the system [Self 94].

In representation of an ability, there are not only numeric values. Other informations are necessary to represent an ability such as those related to the designation of the ability, the unit of measurement, the target knowledges, and the advices. It is necessary to structure these informations. The structuration of data is common to all abilities. It could therefore be pertinent to use prototypes that one could instancied, on the one hand for every ability while bringing in the relative data to every ability, and on the other hand for every student while bringing in values of his or its productions.

The last characteristic of this knowledge is that they are not each independent of the others. For example, if one wants to pronounce a vowel in a steady and intelligible manner, one must have a steady fundamental frequency. Therefore, it is not only necessary to organize knowledge of each parameter with regard to the others, but also to propose a representation that expresses relations of dependence and proximity between abilities. The system of chosen representation must preserve relations of mutual dependence or the proximity of mechanisms between abilities.

We have thus far presented the characteristics and the essential requirements of manipulated knowledge in this system. In the following section, we expose the styles of retained representation.

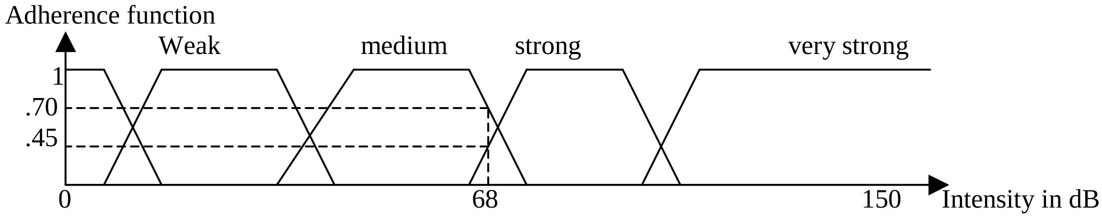
Representation mechanisms in GEREV

From the quantitative to the qualitative

For the speech-therapist, domain's knowledge is expressed in the following form:"the intensity must remain at a medium level". In GEREV, we have chosen to represent speech abilities in as close a form (actually identical if possible) to those manipulated by the specialist. That choice must permits us to transform the quantitative to qualitative while maintaining distances between the numeric values. Therefore, it appeared that fuzzy qualifications would be the most suitable for this type of representation. Thus, for each acoustic parameter, one defines some fuzzy labels such as "weak", "normal", "strong".

For example, for the parameter intensity, several fuzzy labels are defined: "very weak intensity", "weak intensity", "medium intensity", "strong intensity" and "very strong intensity". In Figure 1, we propose an example of that which could be the fuzzy sets and the distributions for the intensity parameter. An intensity of 68 dB would be categorized with a degree of adherence of .70 in "medium intensity" and .45 in "strong intensity".

Intensity frame
Ability moderated with dB
Value of category (between 10 and 150)



value of production : (waited value)
target category : medium intensity
if value of production is different to target category : reeducation in function to the value of production + advices adapted

Parameter’s categories constructed with the help of these fuzzy sets constitute the reference knowledge with wich the system works. The use of fuzzy sets allows to represent speech abilities manipulated by the learner in the same way as the specialist, and to benefit to the type of fuzzy set categorization. This last point is particularly important for the reliability of the datum representation. By this type of data qualification, a precise picture is obtained by going to give a degree of adherence of the learner's production; this is the case for all fuzzy sets defined for this ability. One therefore has, in a way, a picture from different perspectives of the same production, which then facilitates this type of representation. Besides, the use of fuzzy sets permits to account for the evolution of ability for all student's answers. Therefore, thanks to the comparison of two productions, the system can say if there is an evolution of the abilities concerning this acoustic parameter.

Representation and organization of an ability

We have chosen to represent student's productions in a qualitative way, thanks to the fuzzy sets. Nevertheless, this only represents one part of the data necessary for modelling an ability. Indeed, the representation of an ability contains both information on the production’s value, its designation and its type, as well as a common organization to all abilities. These two notions (structuring of different information and identical structuring of all abilities) made us choose to use *frames*. The frames formalism could be compared to forms which include slots to fill in and slots filled by default; it represents prototypes of objects [Minsky, 88]. A frame possesses some of essential properties, notably those of values associated to an attribute by default when the real value is unknown, those of constraints that must satisfy attributes, those of procedures that release when the value of the attribute is either required or newly appeared [Haton91].
The idea of prototype is interesting in this system. It intervenes, on the one hand on each ability in bringing in the specific data to each parameter (its unit of measure, the advice, the category targets), and on the other hand on each learner when the system is instanciated with a numeric value transformed thereafter in qualitative terms. As far as this last point, frames allow a system to be used despite the incomplete or current knowledge on the learner (values by default or values in waiting [Minsky, 88]). Indeed, our student profiles are structures of prototypical data in their form and in their content, in which the entry of data is waited (either learner’s production or manual entry by the specialist) in order to particularize the model and

thus, the system. We present in Figure 2 an outline of the minimum informations which a single frame will contain: informations pertaining to the designation of the frame, the reference's knowledge represented by a distribution of fuzzy sets, the knowledge's target which is the category toward which the leraner's productions must be offered, advices attached to the reference's categories, and the procedure to follow in case of deviation of the target category. Other informations will be able to be represented as exercises in which abilities are treated .

Organization of abilities

Having chosen the type of representation of an ability, we are interested in the type of structuring between them. All abilities forms a knowledge system that constitutes the reference's domain by only considering abilities and not their particularization by learners. To form this structured system, it is necessary to call upon a type of representation which allows to organize some knowledge with regard to others, and to express relations of dependence and proximity between them. The use of a frames network is the most adequate means to fulfil these requirements.

We have thus seen the manner to represent all knowledges manipulated, from the numeric entry to the organization of the knowledge system. We are now going to consider the integration of this representation system in the different components of GEREV: expert and student modules.

Proposition of software architecture of GEREV

In the ITS, there typically are four components [Nicaud and al88] :

- the expert's module in which there are reference's knowledges, and that has an active role because it generally contains strategies used by the expert;
- the student's model, named module in the following, wich are representations of the student's knowledges, possibly in addition to other data such as age, sex, incentives, and the learner's level ;
- a pedagogical module in which there are strategies bound to the aims of the training, such as training by trial and error, etc.
- communication interfaces.

During the last ten years, the most studied of these components has been the student's module. If appropriate to the pedagogical module [Self88; Self94], the student's module should allow individualization of the teaching. In the remainder of this exposition, we consider mainly the expert's and student's module, as they are essentially the ones to which is bound the problem of knowledge representations, or of abilities in our case.

The expert's module

In this section, we will consider, respectively, the following two measurements:

- the data attached to the reference knowledge;
- the reference knowledge itself.

With regard to the expert's module, we propose a singular conception; at first, there is no place in which to include the equivalent of problem resolution strategies because, in the process of speech production, there do not exist any conscious strategies on behalf of the speaker, nor of the expert. Otherwise, we attach advices to the parameter's categories in this module and not in the pedagogical module. The explanation of this last point is quite simple: the advice depends on the type of answers produced, and not on the teaching used or on the subject himself. Indeed, if the intensity of the resonant production is too weak, it is necessary to give advice that will enable the subject to compensate for this deficiency, such as "take a deep breath before beginning...", and not the fact that we are in the presence of such a subject rather than another one, or that we use such an pedagogical strategy rather than another one.

In GEREV, what will vary from one subject to another will be, on the one hand, the choice of one type of help presentation which is bound to the rehabilitation or the order in which the specialist will decide to illustrate the rehabilitation (for example, for the exercise on vowel articulation, system can chooses to propose a cut of the vocal tract schematizing the articulation of the vowel targets, or a representation of the delay between the target production and that of the subject); and on the other hand, the degree of requirement that will be selected to judge the adherence of the production to a fuzzy set defined for this ability... Therefore, all learner's answers that will be categorized as belonging to the same knowledge, will receive the same advice.

The knowledges manipulated by the expert can take two forms: target knowledges which are performance's levels to be ideally attained (the norm), and reference abilities which are the basis from which the system is going to be able to treat the student's productions. This translates itself in the expert's module by a marking of the fuzzy set which must be reached (medium intensity) for seen as being the target, as well as by the other sets bound to the ability in question. The learner succeeds when he or she produces a answer which is categorized in the target fuzzy set as having a high enough degree of adherence to this fuzzy set. Use to representation of abilities with the fuzzy sets, we obtain a representation as the whole of the knowledge domain. Consequently, all of the potential answers are modelised in this module, that eather the just answers or also divergent answers, and this fastly and easily implantable.

The student's module

This module is, in a sense, an instance of the expert's module. Indeed, the ideal objective for the learner is to reach the expert performance. The initial frames network corresponds to the a theoretical student model. The entry of individual data such as the learner's identity, his or her sex, and his or her expertise in the various abilities, lead to the initialization of this user's profile. Theoretically, our conception of the system is to approach a model by overlay [Nicaud and al88; Vanlehn88]. Effectively, the system's objective is to bring students to the norm or the abilities conceptualized by the expert. Nevertheless, this module is not a subset of the expert's model since that corresponds to the acquisition of a norm and not of an set of knowledges or an ability. The expert is therefore not considered here as a "super-user" who would master a vast domain of knowledge and who would have one particular type of utilization (at the level of strategy, for example); he is rather a reference, towards which it is necessary to aim the level of performance.

The main difficulty encountered in overlay models is the problem of forming the knowledge domain. Classically, the expert's knowledges are represented as a set of knowledges that are going to compare the student's knowledge. All student's answers that are not present in the expert's knowledges, are considered false. In this conception, acquirement and hiatuse learner are represented. In order to take into account learner's deviations and therefore to allow the system to treat this type of information, system's conceptualizer build up catalogs of mistakes. This construction is expensive in time and does not necessarily foresee the entirety of potential answers. In our case, the use of fuzzy sets allows us to cover the whole of the learner's answers without previously having to build up mistakes catalogs, and without the risk to forgot a case or a combination of cases.

Conclusion

After a summary of the choices of representing abilities in this system and their principal advantages, we propose, once again, an outline of reflections on the potential generalization and limits of our work. Our choices of representation are as follows:

- an ability is qualified by fuzzy sets (the qualification of data allows to preserve the same type of knowledge representation as that manipulated by the specialist ; to

qualify a production progressively, thanks to a adherence's degree to several fuzzy sets at the same time; and to take into account their evolution);
- an ability is an elementary data set organized in a frame (to propose an data organization in which there is a skeleton of informations with data by default and data necessary to be entered into the system);
- the ability sets or the elementary frames, that either is at the expert's module or at the student's module, is organized inside a frames network (to inform on the proximity and the relations amongst the different abilities or knowledges, and both in structuring the knowledge system).

Now, the problem is to know the extent to which this combination of types of representation can be generalised to other systems. In order to present a clear exposition, we will successively consider the different aspects of types of representation.

The use of the fuzzy qualification permits us to easily represent the entirety of the potential student's answers. This is not a insignifiante contribution, as it theoretically allows to suppress the major default of overlay models. Nevertheless, it is not the fuzzy logic that solves this problem. Indeed, that is possible to the extent that the domain knowledge is given by imprecise data (measures of parameters) that will be categorized in qualitative terms (cover all values consistent of the abilities). The knowledge used must therefore be in concordance with the fuzzy representation type in order for take full advantage of the latter. Nevertheless, it may be judicious to qualify student's answers by fuzzy sets of the type "very close to the correct answer". This might allow to construct a student's model that would take into account not only the student's answer, but also the gap with respect to the target answers. However, the conception of these sets would be relatively expensive, as it would first be necessary to conduct a study in order to foresee all cases and also to appraise their gaps with the target.

The data's organization with frames is very interesting in EIAO, wichever the nature of the knowledge representation. Effectively, it allows to construct the student's model in bringing in default values. On the one hand, this permits the system to function despite the incompleteness of the model; and on the other hand, those values will be subsequently re-estimated in order to work out the student's profile, and in also allowing to propose a knowledge structuration that foresees the new future data. This last type of organization - the network of frames- is quite applicable when the information that the system is to manipulate embodies the notions of dependence or proximity between concepts or knowledges. These notions are independent of the type of information itself, but allow to add and to organize the relative data either to the domain of application or to the student. They therefore appear to be fundamental in the development of the system.

Perspectives

To conclude this exposition, we will succinctly present problems of implementation of GEREV, and student's assessment, reeducation and learner's guidance in exercices (he specifications are finished, the interface are realised and I'm coding the Gerev's system).

The difficulty encountered in implementation is located, on the one hand, at the information exchanges between GEREV and SIRENE, and, on the other hand, at the conception to reference sets (fuzzy sets). For the latter, it is necessary to establish the limits of each set, with the help of the speech-therapist. It is this setting up of limits that will partly depend the exactness of the subject's answer categorization. Nevertheless, while this stage is important, it does not appear insoluble or particularly expensive in the process of development. As for exchanges of informations between GEREV and SIRENE, we foresee the use of an architecture of the " épiphyte" type [Pachet96], which should allow the two systems to function independently. This type of architecture is going to allow GEREV to constantly take information on the subject's productions from SIRENE without interrupting the latter; and

also to play the role of the speech-therapist in sending back data resulting from the assessment, which will allow to guide the subject in the course of the different exercises.
Several times in this article, we have evoked the adaption of the reference knowledge to the learner. We foresee that this possibility can be achieved by a displacement of the limits of fuzzy sets. If a student presents a very weak intensity, these fuzzy sets will displace in order having "normal intensity" between 40 and 60 dB rather than between 50 and 70 dB. This example is theorical since no experimentation or study has currently been carried out to allow adjustment of reference knowledge to a given subject, especially at a given moment. There are different levels of requirement, since one answer can be accept as normal in spite of the relatively distant productions of "medium parameter" sets for subjects having an important deficit in the given parameter.
As for the reeducation and the student's guidance, two types of information are taken into account:

- data on the subject's performance on a parameter;
- prerequisites to the rehabilitation of parameters.

We will realise a work for rehabilitate a parameter when the reference category for a subject is distinct relative to the target category. The student guidance during exercises which make up the rehabilitative progression, is based on prerequisites for each exercise. To perform an exercise, it will be necessary that all parameters relevant to the realization of the exercise are either categorized in "medium parameter" sets or chosen deliberately by the learner or the specialist.
Finally, we foresee two types of subject's assessment: an assessment in context, which corresponds to the assessment of an ability in an exercise specifically designed to assess that ability; and an assessment out of context, which corresponds to an assessment of an ability in certain exercises not directly aimed at the rehabilitation of that ability, but in which that ability plays a part.